Homework 8

Christophe Hunt

March 25, 2017

Contents

1 Problem Set 1

Your colleague either commutes by train or by the bus. 20 days of the month, she takes the train and the remaining 10 days she takes the bus. If she takes the train, she reaches work on time with a probability of 0.9. If she takes the bus, she frequently gets stuck in traffic and reaches work on time with a probability of 0.5. Given that she was on time today, what is the probability that she took the bus to work today?

$$P(bus|on\ time) = \frac{P(on\ time|bus)*P(Bus)}{P(on\ time|bus)*P(Bus)+P(on\ time|Train)*P(Train)}$$

```
library(scales)
train <- 20/30
bus <- 10/30
ontime_train <- .9
ontime_bus <- .5

bus_ontime <- (ontime_bus * bus)/ ((ontime_bus * bus) + (ontime_train*train))
percent(bus_ontime)
```

[1] "21.7%"

The probability that my colleague took the bus given that she is on time is 21.7%

2 Problem Set 2

In the Grade Network that we looked at in the notes,

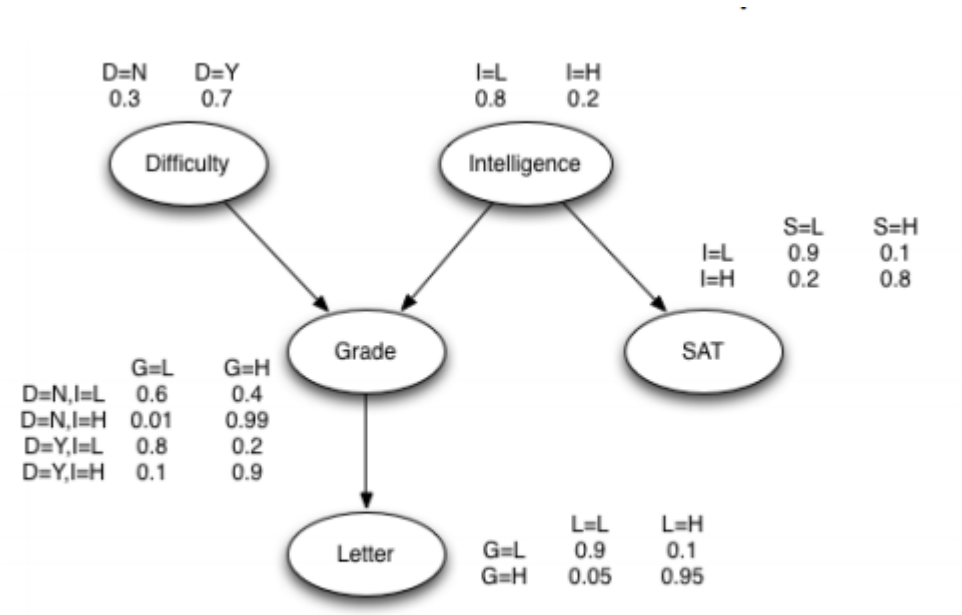

FIGURE 1. A Simple Bayesian Network

Figure 1:

```
library(gRain)
# Build the network
n.y <- c("no", "yes")
l.h <- c("low", "high")
d <- cptable(~difficulty, values = c(0.3, 0.7), levels = n.y)
i <- cptable(~intelligence, values = c(0.8, 0.2), levels = l.h)
s.i <- cptable(~sat | intelligence, values = c(0.9, 0.1, 0.2, 0.8), levels = l.h)
g.d.i <- cptable(~grade | difficulty:intelligence, values = c(0.6, 0.4, 0.8,
    0.2, 0.1, 0.99, 0.1, 0.9), levels = l.h)
l.g <- cptable(~letter | grade, values = c(0.9, 0.1, 0.5, 0.95), levels = l.h)
p <- compileCPT(list(d, i, s.i, g.d.i, l.g))
pn <- grain(p)
```

2.1 What happens to the probability of Difficulty of Course when you present the evidence that the received recommendation letter was good?

```
l.g <- setFinding(pn, nodes = c("letter"), states = c("high"))
kable(as.data.frame(querygrain(l.g, nodes = c("difficulty"), type = "marginal")))
```

	difficulty
no	0.3597003
yes	0.6402997

2.2 In addition, now present the evidence that both SAT scores were good and the letter of recommendation was good, What is the probability of the Difficulty of Course now?

```
s.l.g <- setFinding(setFinding(pn, nodes = c("letter"), states = c("high")),
    nodes = c("sat"), states = c("high"))
kable(as.data.frame(querygrain(s.l.g, nodes = c("difficulty"), type = "marginal")))
```

	difficulty
no	0.3174519
yes	0.6825481

You should use the gRain package in R to build your network and perform these calculations.

You may need to install RBGL package from BioConductor in R to get gRain working. See http://www.bioconductor.org/packages/release/bioc/html/RBGL.html for instructions on RBGL.

Please submit your assignment as an R markdown document.